

Natural risk measures

Hirbod Assa^{*†}

Abstract

A coherent risk measure with a proper continuity condition cannot be defined on a large set of random variables. However, if one relaxes the sub-additivity condition and replaces it with co-monotone sub-additivity, the proper domain of risk measures can contain the set of all random variables. In this study, by replacing the sub-additivity axiom of law invariant coherent risk measures with co-monotone sub-additivity, we introduce the class of natural risk measures on the space of all bounded-below random variables. We characterize the class of natural risk measures by providing a dual representation of its members.

1 Introduction

Although Value at Risk (VaR) is the most popular risk measure among practitioners, it has been heavily criticized in the theoretical literature since it does not necessarily associate portfolio diversification with risk reduction. Therefore, axiomatically founded risk measures such as coherent (cf. Artzner et al. (1999) for finite and Delbaen (2002) for infinite spaces) and, more generally, convex (cf., Föllmer and Schied (2002) and Frittelli and Rosazza Gianin (2002)) risk measures have been introduced whose axioms of sub-additivity, respectively convexity, directly ensure that diversification reduces the measured risk.

This paper argues that all types of risk measures that are studied in the literature suffer from either of the following two drawbacks: either, like coherent and convex risk measures, they cannot be defined on the space of all random variables (see discussions

^{*}Institute for Financial and Actuarial Mathematics, University of Liverpool.

[†]The author would like to sincerely thank the reviewers and in particular the associate editor, Alexander Zimper, for their valuable comments, which considerably improved the quality of this paper.

in Delbaen (2002)) or, like VaR, they might not be available for assessing the risk that arises from model uncertainty (see example 3 below). To address both shortcomings, we are going to introduce and axiomatically characterize the class of natural risk measures. In contrast to the sub-additivity axiom of coherent risk measures, natural risk measures require sub-additivity only for co-monotonic random variables. Note that VaR, first, satisfies additivity for co-monotonic random variables and is, second, well-defined on the space of all random variables. To relax sub-additivity in favor of co-monotonic sub-additivity therefore amounts to studying a general class of risk measures that contains VaR and is well-defined on large spaces of random variables.

The representation of natural risk measures, provided in Theorem 1, can be regarded as a convexification of the representation of insurance risk premiums in Wang et al. (1997)¹. However, we will see that this convexification is not mathematically straightforward and needs a significant amount of further mathematical work. Whereas Wang et al. (1997) used results attributed to Greco in Denneberg (1994) (see Greco (1982) in Italian) to provide the dual representation of insurance risk premiums, we use Daniell integrals to extend the representation of *natural risk statistics* (see Kou et al. (2013) and Ahmed et al. (2008)) on finite probability spaces to the set of all bounded-below random variables. In addition, our representation in Theorem 1 sheds some light on the axiomatic foundations of the VaR criterion versus the Kusuoka (2001) representation, and provides further tools to develop theories of risk measures and risk premiums for large spaces.

The paper is organized as follows: in Section 2, we provide some preliminary mathematical definitions and introduce natural risk measures and the weak continuity. In Section 3, we present examples of natural risk measures that are different from VaR but are also co-monotone sub-additive (and not necessarily additive) in the presence of model uncertainty. In Section 4, we state our main result for Theorem 1, which gives a dual representation of weakly continuous natural risk measures.

2 Preliminaries and definitions

In this section, we will introduce preliminary mathematical tools and definitions with the necessary economic and financial concepts that we will use in our discussions.

¹This convexification can only be regarded as a technical extension since in Wang et al. (1997) the authors have a different objective: to compare their co-monotone additive premium function in a competitive market with an arbitrage-free pricing rule, where additivity holds for all risks.

2.1 Mathematical framework

Let (Ω, \mathcal{F}, P) be an atom-less probability space, where Ω represents the “states of the nature”, \mathcal{F} is the sigma-field of all measurable sets, and P is the physical probability measure. In this study, we consider that L^0 , the set of all measurable functions or random variables on (Ω, \mathcal{F}, P) , represents the set of individual loss variables². Let us also denote the set of all bounded-below random variables by L_B^0 . The space L^0 is a metric space whose metric is defined as $d(X, Y) = E(\min\{|X - Y|, 1\})$, where E denotes the expectation. Convergence in this topology is equivalent to convergence in probability, i.e., $d(X_n, X) \rightarrow 0$ iff $\forall \epsilon > 0, P(|X_n - X| > \epsilon) \rightarrow 0$. The space L^p , for $0 < p$, is the space of all random variables with p -th finite moment, i.e., $L^p = \{X \in L^0 | E(|X|^p) < \infty\}$. L^∞ is the set of all almost surely bounded members of L^0 .

The cumulative distribution function of a random variable $X \in L^0$ is denoted by F_X . For any $X \in L^0$, F_X is a càdlàg³ and non-decreasing function, with a left inverse given by $F_X^{-1}(\alpha) = \inf\{x \in \mathbb{R} : F_X(x) \geq \alpha\}$, for $\alpha \in (0, 1)$, which is also a càdlàg function. If $X \in L_B^0$, one can extend the inverse to $\alpha = 0$, i.e., $F_X^{-1}(0) = \text{essinf}(X)$, where essinf is the essential infimum.

Two random variables $X, X' \in L^0$ have the same distribution if and only if $F_X = F_{X'}$. Two random variables, X and Y , are co-monotone if

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0 \quad \text{a.s. } \omega, \omega' \in \Omega.$$

In this study, we use a version of co-monotonicity due to Denneberg (cf. Proposition 2 in Wang et al. (1997)), which says that X and Y are co-monotone if there are two non-decreasing real functions f and g and a random variable U such that $X = f(U)$ and $Y = g(U)$. Finally, as usual, $\mathcal{B}[0, 1]$ denotes the set of all Borel measurable subsets of $[0, 1]$.

2.2 Natural risk measures

Now we introduce the class of natural risk measures

Definition 1. A natural risk measure⁴ ϱ is a mapping from L_B^0 to \mathbb{R} that satisfies the following conditions:

²Unlike in financial mathematics literature, which considers a profit variable, we found the loss variable more convenient to deal with.

³Càdlàg is a French acronym that translates into English as “right continuous and left limited.”

⁴The definition of a *natural risk measure* is motivated by the definition of a *natural risk statistics* introduced on \mathbb{R}^n in Kou et al. (2013).

1. Positive homogeneity: $\varrho(\lambda X) = \lambda \varrho(X), \forall \lambda > 0$ and $\forall X \in L_B^0$;
2. Cash invariance: $\varrho(X + c) = \varrho(X) + c, \forall X \in L_B^0$ and $\forall c \in \mathbb{R}$;
3. Monotonicity: $\varrho(X) \leq \varrho(Y), \forall X, Y \in L_B^0$ and $X \leq Y$;
4. Co-monotone sub-additivity: $\varrho(X + Y) \leq \varrho(X) + \varrho(Y), \forall X, Y \in L_B^0$; X and Y are co-monotone;
5. Law invariance: $\varrho(X) = \varrho(Y)$ if $F_X = F_Y$, i.e., X and Y have the same distribution.

Let

$$\mathbb{F} = \left\{ K : [0, 1] \rightarrow \mathbb{R} \cup \{+\infty\} \left| \begin{array}{l} K(0) \in \mathbb{R} \\ K(1) \in \mathbb{R} \cup \{+\infty\} \\ \forall \alpha \in (0, 1), K(\alpha) \in \mathbb{R} \end{array} \right. \right\}$$

and $\mathbb{A} = \mathbb{F} \cap \{\text{non-decreasing and càdlàg functions}\}$. It is clear that \mathbb{A} consists of all left inverse cumulative distribution functions with a finite essential infimum. It is known that, for any random variable X and any random variable U with a uniform distribution on $[0, 1]$, X and $F_X^{-1}(U)$ have the same distribution. That is why we can consider a natural risk measure ϱ as a well-defined function on the set of all inverse cumulative distribution functions \mathbb{A} . We use this fact later in the proof of our main result for Theorem 1.

Definition 2. The natural risk measure ϱ is weakly continuous if $\varrho(X_n) \xrightarrow{n \rightarrow \infty} \varrho(X)$ when $F_{X_n}(x) \xrightarrow{n \rightarrow \infty} F_X(x), \forall x \in \mathbb{R}$.

3 Examples of natural risk measures

In this section, we introduce co-monotone additive and sub-additive natural risk measures that naturally emerge in insurance and finance applications.

Example 1. Value at Risk, for a fixed tolerance level $\alpha \in [0, 1)$, is introduced as follows:

$$\text{VaR}_\alpha(X) = F_X^{-1}(\alpha), X \in L_B^0.$$

Note that, since natural risk measures are defined on L_B^0 , $\text{VaR}_0(X) = F_X^{-1}(0) = \text{essinf}(X) \in (-\infty, +\infty)$.

Observe that VaR is weakly continuous. To see this, let us consider $F_{X_n}(x) \xrightarrow[n \rightarrow \infty]{} F_X(x), \forall x \in \mathbb{R}$. First, we prove that there is no $\epsilon > 0$ such that $\text{VaR}_\alpha(X_n) \leq \text{VaR}_\alpha(X) - \epsilon, \forall n$. Indeed, if this happens, then for any y such that $F_X(y) \geq \alpha$ we have that $\text{VaR}_\alpha(X_n) \leq y - \epsilon, \forall n$. By right continuity and monotonicity of F_{X_n} , this implies that $\alpha \leq F_{X_n}(y - \epsilon)$. When n tends to infinity, then this gives that $\alpha \leq F_X(y - \epsilon)$. This implies that $\text{VaR}_\alpha(X) = \inf \{x \in \mathbb{R} : F_X(x) \geq \alpha\} \leq y - \epsilon$. Now, by taking infimum over all y such that $F_X(y) \geq \alpha$, we get that $\text{VaR}_\alpha(X) \leq \text{VaR}_\alpha(X) - \epsilon$, which is a contradiction. Now, using the fact that $\text{VaR}_\alpha(X) = -\text{VaR}_{1-\alpha}(-X)$, one can show that there is no $\epsilon > 0$ such that $\text{VaR}_\alpha(X_n) \geq \text{VaR}_\alpha(X) + \epsilon, \forall n$.

Example 2. Consider an insurance company whose total loss for a fiscal year can be represented by a non-negative random variable X . Consider also that the insurance company will have to buy a reinsurance contract $0 \leq Y \leq X$. The premium of Y is simply given by expectation, i.e., $E(Y)$. Therefore, the insurance company's global position is $X - Y + E(Y)$. On the other hand, to avoid the risk of moral hazard, the contract Y should be such that both parties feel any increase in the losses. Therefore, we consider a non-decreasing and non-negative function f , so that $x \mapsto x - f(x)$ is also non-decreasing and non-negative; see, for example, Cai and Tan (2007) and Cai et al. (2008) and Assa (2015a). Let us denote the set of all such functions f by \mathbb{C} . If the insurance company measures its risk by VaR_α , for some $\alpha \in (0, 1)$, the optimal contract f^* will be found by solving

$$\min_{f \in \mathbb{C}} \text{VaR}_\alpha((X - f(X)) + E(f(X))).$$

Therefore, the risk of the global position is given by

$$\varrho(X) = \text{VaR}_\alpha(X - f^*(X)) + E(f^*(X)) = \min_{f \in \mathbb{C}} \text{VaR}_\alpha((X - f(X)) + E(f(X))).$$

Following the discussions in Assa (2015a) and Assa and Karai (2013), one can easily see that

$$\varrho(X) = \min_{f \in \mathbb{C}} \text{VaR}_\alpha(X - f(X)) + E(f(X)) = \int_0^1 \text{VaR}_t(X) d\lambda(t),$$

where

$$\lambda(x) = \max \{1_{[\alpha, 1]}(x), x\} = \begin{cases} x & , \alpha < x \leq 1 \\ 0 & , 0 \leq x \leq \alpha \end{cases}.$$

Therefore, to assess the risk of the insurance company's global position, one needs to define a new risk measure as

$$\varrho(X) = (1 - \alpha) \text{VaR}_\alpha(X) + \int_0^\alpha \text{VaR}_t(X) dt,$$

which is different from VaR_α . Note that this risk measure can be defined on the set of all bounded-below random variables.

By abusing the notation, one can consider λ as a measure on $[0, 1]$ defined as $\lambda(a, b] = \lambda(b) - \lambda(a)$. Observe that the support of λ as a measure is $[0, \alpha)$. According to Theorem 1 in Section 4, this condition implies the weak continuity of the risk measure ϱ .

The following simple example shows how in the presence of model uncertainty, when a robust analysis approach needs to be conducted, one may have a co-monotone sub-additive risk measure⁵.

Example 3. Let us consider an insurance company that needs to issue a deposit insurance on an asset value that follows a geometric Brownian motion dynamics

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dW_t \\ S_0 > 0 \end{cases}, 0 < t \leq T.$$

Here $\{W_t\}_{t=0}^T$ is a standard Brownian motion on $[0, T]$, $\mu \in \mathbb{R}$ is the drift, and $\sigma > 0$ is the volatility. After solving this stochastic differential equation, we get that $S_t = S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right)$, $0 \leq t \leq T$. Let us assume that the losses of the financial company (which needs to be insured) can be given by $X = \mathcal{L}(S_T)$, where \mathcal{L} is a non-increasing real function (e.g., $\mathcal{L}(x) = \max\{e^{rT}S_0 - x, 0\}$, where $r > 0$ is the risk-free rate). As discussed before, in order to avoid the risk of moral hazard, one needs to consider a contract like $f(L)$, where $f \in \mathbb{C}$, implying that a contract Y is a non-increasing function of S_T . Now the market premium of the contract Y can be found as its market value (since it is a European option) given by the discounted expectation under the risk neutral probability:

$$\pi(Y) := e^{-rT} E(\varphi Y),$$

where

$$\varphi = \exp\left(\left(\frac{1}{2}m^2/\sigma^2 - m/2\right)T\right) \left(\frac{\exp(-rT)S_T}{S_0}\right)^{-m/\sigma^2}.$$

Here we have $m = \mu - r$. For instance, see Nakano (2004) for further details. As one can see, if $0 < m < \sigma^2$, then φ is also a non-increasing function of S_T . Following

⁵Robust optimization is an approach to model uncertainty when the uncertain parameters are known to be within certain bounds. For more reading on the robust analysis approach, see Bertsimas et al. (2011), Quaranta and Zaffaroni (2008), Xing et al. (2014), and Zymler et al. (2011).

Assa (2015b), one can see that

$$E(\varphi Y) = \int_0^1 \text{VaR}_t(\varphi) \text{VaR}_t(Y) dt = \int_0^1 \text{VaR}_t(Y) d\lambda_1(t),$$

where $\lambda_1(x) = \int_0^x \text{VaR}_t(\varphi) dt$. One can show further that $\lambda_1(x) = N\left(N^{-1}(x) - \frac{m\sqrt{T}}{\sigma}\right)$, where $N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$ is the cumulative distribution function of the standard normal distribution. Similar to the previous example, the risk of the global position will be assessed by a risk measure ϱ given by

$$\varrho(X) = \int_0^1 \text{VaR}_t(X) d\lambda(t), \quad (1)$$

where $\lambda(x) = \max\{1_{[\alpha,1]}(x), \lambda_1(x)\}$.

Now, let us consider that there is uncertainty in estimating the volatility σ . That means that, for two positive numbers σ_{\min} and σ_{\max} , we only know $\sigma \in [\sigma_{\min}, \sigma_{\max}]$, where $0 < m < \sigma_{\min}^2$. In that case, the risk has to be assessed in a robust manner

$$\begin{aligned} \varrho^{\text{Robust}}(X) &= \sup_{\sigma_{\min} \leq \sigma \leq \sigma_{\max}} \int_0^1 \text{VaR}_t(X) d\lambda^\sigma(t) \\ &= (1 - \alpha) \text{VaR}_\alpha(X) + \int_0^\alpha \text{VaR}_t(X) d\left(N\left(N^{-1}(t) - \frac{m\sqrt{T}}{\sigma}\right)\right) \\ &= (1 - \alpha) \text{VaR}_\alpha(X) + \sup_{\sigma_{\min} \leq \sigma \leq \sigma_{\max}} e^{\frac{-m^2 T}{2\sigma^2}} \int_0^\alpha \text{VaR}_t(X) e^{\frac{m\sqrt{T}N^{-1}(t)}{\sigma}} dt, \end{aligned}$$

where λ^σ is from (1). As one can see, ϱ^{Robust} is co-monotone sub-additive but not necessarily co-monotone additive. Note that this risk measure can be defined on the set of all bounded-below random variables. With a similar argument as in the previous example, if we look at λ^σ as a measure, the support of all measures in $\{\lambda^\sigma\}_{\sigma \in [\sigma_{\min}, \sigma_{\max}]}$ is $[0, \alpha]$. Again, according to Theorem 1 in Section 4, this shows that ϱ^{Robust} is weakly continuous.

4 Dual characterization of natural risk measures

In this section, we characterize the family of weak continuous natural risk measures. However, in order to present our main result, we need to introduce some further notations.

Let $C[0, 1]$ be the space of all continuous functions on $[0, 1]$ with the uniform norm $\|\cdot\|_\infty$. Then it is known that the topological dual of $C[0, 1]$ is the space of all bounded variation functions on $[0, 1]$, denoted by $BV[0, 1]$, with the total variation on $[0, 1]$ as its norm. The dual relation between $C[0, 1]$ and $BV[0, 1]$ is defined as

$$\langle H, K \rangle = \int_0^1 K(t) dH(t), \forall (K, H) \in C[0, 1] \times BV[0, 1],$$

where the integral is the Riemann-Stieltjes integral. In the following discussions, $\langle H, K \rangle$ is used to show $\int_0^1 K(t) dH(t)$ when in general K is H integrable. The same bi-linear operator introduces the smallest topology on $BV[0, 1]$, with the topological dual $C[0, 1]$. This topology is denoted by $\sigma(BV[0, 1], C[0, 1])$ ⁶.

Now we are in a position to state the main result of this study.

Theorem 1. *Let $\varrho : L_B^0 \rightarrow \mathbb{R}$ be a natural risk measure characterized by Definition 1. Then ϱ is weakly continuous if and only if there exists a compact set Δ in $\sigma(BV[0, 1], C[0, 1])$ such that*

1. *Each $\lambda \in \Delta$ is a probability measure on $([0, 1], \mathcal{B}[0, 1])$;*
2. *There exists $\epsilon_0 > 0$ such that $\forall \lambda \in \Delta$, $\text{supp}(\lambda) = [0, 1 - \epsilon_0]$* ⁷

and

$$\varrho(X) = \sup_{\lambda \in \Delta} \int_0^1 \text{VaR}_t(X) d\lambda(t). \quad (2)$$

Moreover, if ϱ is co-monotone additive, then $\Delta = \{\lambda\}$.

Before presenting the proof of the theorem, we need to introduce some notations and recall some statements from the convex analysis that we will use in the proof.

Let V be a topological vector space, and let V' be its topological dual. Recall the Alaoglu theorem that states that for any set $C \subseteq V$, with a non-empty $\sigma(V, V')$ -interior and any $c \in \mathbb{R}_+$, the set $\{H \in V' \mid \sup_{K \in C} \langle H, K \rangle \leq c\}$ is $\sigma(V', V)$ compact. Also, recall that V has the Dunford-Pettis property if for any sequence $\{(F_n, H_n)\}_{n=1}^\infty \subseteq V \times V'$ converging in $\sigma(V, V') \times \sigma(V', V)$ to (F, H) , $\{\langle F_n, H_n \rangle\}_{n=1}^\infty$ converges to $\langle F, H \rangle$. It is known that $C[0, 1]$ has this property.

Let us assume that V is a locally convex topological vector space. Recall that the domain of any convex function $\phi : V \rightarrow \mathbb{R} \cup \{+\infty\}$, denoted by $\text{dom}(\phi)$, is equal to $\{K \in V \mid \phi(K) < \infty\}$. The dual of ϕ , denoted by ϕ^* , is defined as $\phi^*(K) =$

⁶In general, for any two topological vector spaces V, V' with bi-linear dual relation (v, v') , $\sigma(V, V')$ denotes the smallest topology on V under which all members of V' are continuous.

⁷ $\text{supp}(\lambda)$ stands for support of λ .

$\sup_{H \in V'} \{\langle H, K \rangle - \phi(H)\}$. A convex function is said to be lower semicontinuous iff $\phi = \phi^{**}$. For a closed convex set $C \subseteq V$, the indicator function of C , denoted by χ_C , is introduced as $\chi_C(X) = 0$ if $X \in C$ and $+\infty$ otherwise. For any positive homogeneous convex function ϕ , let

$$\Delta_\phi = \{H \in V' \mid \langle H, K \rangle \leq \phi(K), \forall K \in V\}.$$

It is easy to see that $\phi^* = \chi_{\Delta_\phi}$. Therefore, any positive homogeneous function ϕ can be represented as $\phi(K) = \sup_{H \in \Delta_\phi} \langle H, K \rangle$. If ϕ is continuous, then by the Alaoglu theorem we know that Δ_ϕ is $\sigma(V', V)$ -compact.

Remark 1. For the reader's benefit, we recommend Grothendieck (1973) for the functional analysis discussions. This book contains all the tools and definitions that are used in this paper, in particular the Alaoglu and the Dunford-Pettis theorems. One can also read Rudin (1987) and Rudin (1991) for further information. For the convex analysis part, we also recommend to see Ekeland and Témam (1999) and Rockafellar (1997).

Proof of Theorem 1. First, we prove the simple implication that ϱ in (2) is a weakly continuous natural risk measure. We leave it to the reader to check that ϱ satisfies the condition of a natural risk measure. We only check that any functional, in the form of (2), is weakly continuous.

Given that for every $X \in L^0$, $\text{VaR}_t(X)$ is a non-decreasing function of t , ϱ is finite on L_+^0 . Now, let us consider that $X_n \geq 0$ converges weakly to X , as n tends to infinity. This implies that $\text{VaR}_t(X_n)$ converges pointwise to $\text{VaR}_t(X)$ for $t < 1$. Let $N \in \mathbb{N}$ be large enough such that $\text{VaR}_{1-\epsilon_0}(X_n) \leq \text{VaR}_{1-\epsilon_0}(X) + 1$, for all $n \geq N$. Since $\text{VaR}_t(\cdot)$ is non-decreasing in t , this implies that $0 \leq \text{VaR}_t(X_n) \leq \text{VaR}_{1-\epsilon_0}(X) + 1, t \leq 1 - \epsilon_0$, for all $n \geq N$. Let $\{K_n\}_{n=1}^\infty$ be a sequence of continuous functions such that $|\text{VaR}_t(X_n) - \text{VaR}_t(X)| \leq K_n(t), t \in [0, 1 - \epsilon_0], n = N, N + 1, \dots$, and $K_n \rightarrow 0$, pointwise. Since $|\text{VaR}_t(X_n) - \text{VaR}_t(X)| \leq 2\text{VaR}_{1-\epsilon_0}(X) + 1$ for $t \in [0, 1 - \epsilon_0]$ and $n = N, N + 1, \dots$, one can consider that $\{K_n\}_{n=N}^\infty$ is bounded above. Therefore, for every $\mu \in BV[0, 1]$, by using the dominated convergence theorem, we have that $\int_0^1 K_n(t) d|\mu|(t) \rightarrow 0$.⁸ This implies that $K_n \rightarrow 0$ in $\sigma(C[0, 1], BV[0, 1])$. Since Δ is $\sigma(BV[0, 1], C[0, 1])$ compact, for each $n \geq 1$ there exists $\lambda_n \in \Delta$ such that $\sup_{\lambda \in \Delta} \int_0^1 K_n(t) d\lambda(t) = \int_0^1 K_n(t) d\lambda_n(t)$.

We prove the continuity of ϱ by way of contradiction. Assume there exist $\delta > 0$ and a sub-sequence $\{X_{n_i}\}_{i=1}^\infty$ such that $|\varrho(X_{n_i}) - \varrho(X)| \geq \delta$. Let $\{\lambda_{n_{i_k}}\}_{k=1}^\infty$ be a

⁸ $|\mu| = \mu_+ + \mu_-$ is the absolute value of μ .

sub-sequence that converges to $\lambda \in \Delta$ in $\sigma(BV[0, 1], C[0, 1])$; then since $C[0, 1]$ has the Dunford-Pettis property, it follows that $\int_0^1 K_{n_{i_k}}(t) d\lambda_{n_{i_k}}(t) \rightarrow 0$ as $k \rightarrow \infty$. Now we have

$$\begin{aligned}
0 < \delta &\leq \left| \varrho(X_{n_{i_k}}) - \varrho(X) \right| \leq \left| \sup_{\lambda \in \Delta} \int_0^1 \text{VaR}_t(X_{n_{i_k}}) d\lambda(t) - \sup_{\lambda \in \Delta} \int_0^1 \text{VaR}_t(X) d\lambda(t) \right| \\
&\leq \sup_{\lambda \in \Delta} \int_0^1 \left| \text{VaR}_t(X_{n_{i_k}}) - \text{VaR}_t(X) \right| d\lambda(t) \\
&\leq \sup_{\lambda \in \Delta} \int_0^1 K_{n_{i_k}}(t) d\lambda(t) \\
&= \int_0^1 K_{n_{i_k}}(t) d\lambda_{n_{i_k}}(t) \rightarrow 0,
\end{aligned}$$

which is a contradiction. This completes the proof of the first implication.

Now we prove the second implication, i.e., we show that, if ϱ is a weakly continuous natural risk measure, then there exists a compact set Δ in $\sigma(BV[0, 1], C[0, 1])$ as described by conditions 1 and 2 in the theorem statement such that $\varrho(X) = \sup_{\lambda \in \Delta} \int_0^1 \text{VaR}_t(X) d\lambda(t)$.

First, we give an outline of the proof as follows:

1. We restrict ϱ to

$$S_n := \left\{ \sum_{i=1}^{2^n} x_i 1_{\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)} + x_{2^n+1} 1_{\{1\}} \mid (x_i)_{i=1}^{2^n+1} \in \mathbb{R}^{2^n+1} \right\}.$$

Then it is clear that $\varrho|_{S_n}$ is a natural risk statistics⁹.

2. Using the previous step, we extend ϱ to $\bigcup_{n=1}^{\infty} S_n$.
3. By using the previous step and the conditional expectation on the partition $\left\{ \left[\frac{i-1}{2^n}, \frac{i}{2^n} \right), i = 1, \dots, 2^n, \{1\} \right\}$, we extend ϱ to $C[0, 1]$.
4. We give the Fenchel-Moreau representation of ϱ on $C[0, 1]$.
5. By using the Daniell integral, we extend the $C[0, 1]$ - Fenchel-Moreau representation of ϱ to \mathbb{A} .

⁹One would wonder why we use the term 'statistics' instead of 'statistic.' Actually, there is no reason except that it is the exact term that has been used in the literature; see, e.g., Ahmed et al. (2008).

Let us fix a uniformly distributed random variable U . We introduce the risk measure Π on \mathbb{A} by

$$\Pi(H) = \varrho(H(U)).$$

Let $H_1, H_2 \in \mathbb{A}$, then it is clear that $H_1(U)$ and $H_2(U)$ are co-monotone. Since ϱ is co-monotone sub-additive, it implies that Π is sub-additive over \mathbb{A} . It is also clear that Π is a positive homogeneous mapping of degree 1 on \mathbb{A} .

Now, let us introduce the following functions from \mathbb{R}^{2^n+1} to \mathbb{A} :

$$T_n(\mathbf{x}) = \sum_{i=1}^{2^n} x_i 1_{\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)} + x_{2^n+1} 1_{\{1\}}, \forall \mathbf{x} = (x_i)_{i=1}^{2^n+1},$$

and

$$\tilde{T}_n(\mathbf{w}) = 2^n \sum_{i=1}^{2^n} w_i 1_{\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)} + w_{2^n+1} 1_{\{1\}}, \forall \mathbf{w} = (w_i)_{i=1}^{2^n+1}.$$

Let $S_n = T_n(\mathbb{R}^{2^n+1})$ and $S = \cup_{n=1}^{\infty} S_n$. We introduce a natural risk statistics Π_n on \mathbb{R}^{2^n+1} , for $n \geq 1$, as follows:

$$\Pi_n(x_1, \dots, x_{2^n+1}) = \Pi(T_n(\mathbf{x}^{os})) = \Pi\left(\sum_{i=1}^{2^n} x_i^{os} 1_{\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)} + x_{2^n+1}^{os} 1_{\{1\}}\right),$$

where $\mathbf{x}^{os} = (x_i^{os})_{i=1}^{2^n+1}$ is the order statistics of $\mathbf{x} = (x_i)_{i=1}^{2^n+1}$, i.e., $x_1^{os} \leq x_2^{os} \leq \dots \leq x_{2^n+1}^{os}$. Define the dual relation between \mathbb{R}^{2^n+1} and itself, with the Euclidean norm, as

$$\langle \mathbf{w}, \mathbf{x} \rangle_n = \sum_{i=1}^{2^n+1} w_i x_i, \forall (\mathbf{w}, \mathbf{x}) \in \mathbb{R}^{2^n+1} \times \mathbb{R}^{2^n+1}.$$

In Ahmed et al. (2008), it is shown that Π_n can be represented as follows:

$$\Pi_n(\mathbf{x}) = \sup_{\mathbf{w} \in \Omega_n} \langle \mathbf{w}, \mathbf{x}^{os} \rangle_n, \forall \mathbf{x} \in \mathbb{R}^{2^n+1},$$

where Ω_n is a closed convex subset of

$$\Lambda_n = \left\{ \mathbf{w} \in \mathbb{R}^{2^n+1} \mid w_i \geq 0, i = 1, \dots, 2^n + 1, \sum_{i=1}^{2^n+1} w_i = 1 \right\}.$$

Let us introduce Γ_n as

$$\Gamma_n(\mathbf{x}) = \sup_{\mathbf{w} \in \Omega_n} \langle \mathbf{w}, \mathbf{x} \rangle_n, \forall \mathbf{x} \in \mathbb{R}^{2^n+1}.$$

Note that, since Γ_n is translation invariant, it is continuous (using the sup norm) on \mathbb{R}^{2^n+1} . Given this, and the discussions we had in the previous section, Ω_n is easily given as follows:

$$\Omega_n = \{ \mathbf{w} \in \mathbb{R}^{2^n+1} \mid \langle \mathbf{w}, \mathbf{x} \rangle_n \leq \Gamma_n(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^{2^n+1} \}.$$

We now introduce the following mapping from \mathbb{R}^{2^n+1} to $\mathbb{R}^{2^{n+1}+1}$:

$$d_n(x_1, x_2, \dots, x_{2^n+1}) = (x_1, x_1, x_2, x_2, \dots, x_{2^n}, x_{2^n}, x_{2^n+1}).$$

Notice that

$$\begin{aligned} \Gamma_n(\mathbf{x}) &= \Pi(T_n(\mathbf{x})) \\ &= \Pi\left(\sum_{i=1}^{2^n} x_i 1_{\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)} + x_{2^n+1} 1_{\{1\}}\right) \\ &= \Pi\left(\sum_{i=1}^{2^n} \left(x_i 1_{\left[\frac{2i-2}{2^{n+1}}, \frac{2i-1}{2^{n+1}}\right)} + x_i 1_{\left[\frac{2i-1}{2^{n+1}}, \frac{2i}{2^{n+1}}\right)}\right) + x_{2^n+1} 1_{\{1\}}\right) \\ &= \Pi(T_{n+1}(x_1, x_1, x_2, x_2, \dots, x_{2^n}, x_{2^n}, x_{2^n+1})) \\ &= \Gamma_{n+1}(d_n(\mathbf{x})), \end{aligned}$$

therefore $\Gamma_n = \Gamma_{n+1} \circ d_n$. Let us define a mapping from $\mathbb{R}^{2^{n+1}+1}$ to \mathbb{R}^{2^n+1}

$$d^n(w_1, \dots, w_{2^{n+1}+1}) = (w_1 + w_2, w_3 + w_4, \dots, w_{2^{n+1}-1} + w_{2^{n+1}}, w_{2^{n+1}+1}).$$

It is very simple to check that

$$\langle \mathbf{w}, d_n(\mathbf{x}) \rangle_{n+1} = \langle d^n(\mathbf{w}), \mathbf{x} \rangle_n, \forall (\mathbf{w}, \mathbf{x}) \in \mathbb{R}^{2^{n+1}+1} \times \mathbb{R}^{2^n+1}.$$

We claim that $d^n(\Omega_{n+1}) = \Omega_n$, and to see this, observe that

$$\begin{aligned} \Gamma_n(\mathbf{x}) &= \Gamma_{n+1}(d_n(\mathbf{x})) \\ &= \sup_{\mathbf{w} \in \Omega_{n+1}} \langle \mathbf{w}, d_n(\mathbf{x}) \rangle_{n+1} \\ &= \sup_{\mathbf{w} \in \Omega_{n+1}} \langle d^n(\mathbf{w}), \mathbf{x} \rangle_n \\ &= \sup_{\mathbf{w} \in d^n(\Omega_{n+1})} \langle \mathbf{w}, \mathbf{x} \rangle_n. \end{aligned}$$

Introduce $\Delta_n = \tilde{T}_n(\Omega_n)$ and

$$L_n(K) := \sup_{H \in \Delta_n} \langle H, K \rangle, K \in L^0,$$

then it is easy to verify that for all $(\mathbf{w}, \mathbf{x}) \in \mathbb{R}^{2^{n+1}+1} \times \mathbb{R}^{2^n+1}$

$$\langle \mathbf{w}, \mathbf{x} \rangle_n = \left\langle \tilde{T}_n(\mathbf{w}), T_n(\mathbf{x}) \right\rangle.$$

This implies that

$$L_n(T_n(\mathbf{x})) = \Gamma_n(\mathbf{x}), \forall \mathbf{x} \in \mathbb{R}^{2^n+1}.$$

Also note that if $\mathbf{x} \in \mathbb{R}^{2^n+1}$ is a non-decreasing sequence, then $T_n(\mathbf{x})$ is non-decreasing and thus

$$L_n(T_n(\mathbf{x})) = \Gamma_n(\mathbf{x}) = \Pi(T_n(\mathbf{x})). \quad (3)$$

Let $T_n(\mathbf{x}) \in S_n$ and $\tilde{T}_{n+1}(\mathbf{w}) \in S_{n+1}$. It can be easily checked that

$$\left\langle \tilde{T}_{n+1}(\mathbf{w}), T_n(\mathbf{x}) \right\rangle = \langle d^n(\mathbf{w}), \mathbf{x} \rangle_n,$$

which implies that

$$\begin{aligned} L_{n+1}(T_n(\mathbf{x})) &= \sup_{H \in \Delta_{n+1}} \langle H, T_n(\mathbf{x}) \rangle \\ &= \sup_{\mathbf{w} \in \Omega_{n+1}} \left\langle \tilde{T}_{n+1}(\mathbf{w}), T_n(\mathbf{x}) \right\rangle \\ &= \sup_{\mathbf{w} \in \Omega_{n+1}} \langle d^n(\mathbf{w}), \mathbf{x} \rangle_n \\ &= \sup_{\mathbf{w} \in \Omega_{n+1}} \langle \mathbf{w}, d_n(\mathbf{x}) \rangle_n \\ &= \Gamma_{n+1} \circ d_n(\mathbf{x}) \\ &= \Gamma_n(\mathbf{x}) \\ &= L_n(T_n(\mathbf{x})). \end{aligned}$$

This shows that

$$L_{n+1}(T_n(\mathbf{x})) = L_n(T_n(\mathbf{x})). \quad (4)$$

Let us introduce the mapping from the set of continuous functions, $C[0, 1]$, to S_n with

$$E_n(K) = \sum_{i=1}^{2^n} \left(\frac{1}{2^n} \int_{\frac{i-1}{2^n}}^{\frac{i}{2^n}} K(t) dt \right) 1_{\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)} + \left(\frac{1}{2^n} \int_{\frac{2^n-1}{2^n}}^1 K(t) dt \right) 1_{\{1\}}, K \in C[0, 1].$$

From a probabilistic point of view, this is the conditional expectation with respect to the sigma-algebra induced by partition $\left\{ \left[\frac{i-1}{2^n}, \frac{i}{2^n}\right), i = 1, \dots, 2^n-1, \left[\frac{2^n-1}{2^n}, 1\right] \right\}$. For

a continuous function K , it is clear that $E_n(K)$ converges pointwise to K . Now, we introduce the following function L^c on $C[0, 1]$

$$L^c(K) = \limsup_n L_n(E_n(K)).$$

First of all, it is clear that $\min(K) \leq L^c(K) \leq \max(K)$; therefore $L^c(K)$ is a finite number, which means $\text{dom}(L^c) = C[0, 1]$. Note that any convex function is continuous in the interior of its domain. Now let us consider a continuous and non-decreasing member $K \in C[0, 1]$; then it is clear that $E_n(K)$ is a non-decreasing member of S_n . Therefore, as discussed earlier, $L(E_n(K)) = \Pi(E_n(K))$. However, since $E_n(K)$ converges pointwise to K , we have that $\Pi(E_n(K)) \rightarrow \Pi(K)$, as $n \rightarrow \infty$. Therefore, by using (4) and (3),

$$L^c(K) = \limsup_n L_n(E_n(K)) = L_n(E_n(K)) = \Pi(K). \quad (5)$$

It can be easily checked that L^c is sub-additive and positive homogeneous of degree 1, that it is non-decreasing, and that $L^c(K + c) = L^c(K) + c$. The last-mentioned property easily results in the continuity of L^c (note $L^c(K) - L^c(H) \leq L^c(H - K) \leq L^c(\|H - K\|_\infty) = \|H - K\|_\infty$, for all $K, H \in C[0, 1]$). Let $C = \{K \in C[0, 1] \mid L^c(K) \geq 0\}$ and

$$\Delta = \left\{ \mu \in BV[0, 1] \mid \forall K \in C, \int_0^1 K(t) d\mu(t) \geq 0 \text{ and } \int_0^1 d\mu(t) = 1 \right\}.$$

Since $C[0, 1]_+ \subseteq C$, then all members of Δ are non-negative and it is also easy to see that Δ is a closed convex set in $BV[0, 1]$. Like in the proof of Theorem 2.3 in Delbaen (2002), one can show that

$$L^c(K) = \sup_{\mu \in \Delta} \int_0^1 K(t) d\mu(t).$$

By using the Alaoglu theorem, we find that the continuity of L^c results in Δ being $\sigma(BV[0, 1], C[0, 1])$ -compact.

We now prove the following lemma:

Lemma 1. *There exists $\epsilon_0 > 0$ such that*

$$\forall \mu \in \Delta, \int_{1-\epsilon_0}^1 d\mu(t) = 0. \quad (6)$$

Proof. Let us consider the opposite to be true and prove it by means of contradictions. Therefore, for any N , there exists μ_N such that $\int_{1-\frac{1}{N}}^1 d\mu_N(t) > 0$. For each N , let us consider a continuous, non-negative, and non-decreasing function I_N such that $\text{supp}(I_N) \subseteq [1 - \frac{1}{N}, 1]$ and $I_N \Big|_{[1-\frac{1}{2N}, 1]} = \frac{1}{\int_{1-\frac{1}{N}}^1 d\mu_N(t)}$. Observe that $I_N \rightarrow 0$, pointwise as $N \rightarrow \infty$, which implies $\Pi(I_N) \rightarrow 0$ as $N \rightarrow \infty$. Since $I_N \in C[0, 1]$ is non-decreasing, then by (5) we have $L^c(I_N) = \Pi(I_N)$. This means that $L^c(I_N) \rightarrow 0$ as $N \rightarrow \infty$. However,

$$\begin{aligned} L^c(I_N) &\geq \int_0^1 I_N(t) d\mu_N(t) \\ &\geq \int_{1-\frac{1}{2N}}^1 \frac{1}{\int_{1-\frac{1}{N}}^1 d\mu_N(t)} d\mu_N(t) = 1, \end{aligned}$$

which is a contradiction. \square

Now, we show that each member $\mu \in \Delta$ can be considered as a measure on $[0, 1]$. First, we have the following simple lemma:

Lemma 2. *For $\mu \in \Delta$, let $\Lambda_\mu(K) = \int_0^1 K(t) d\mu(t)$. Λ has the following properties:*

- 1) *Linearity: If $K_1, K_2 \in C[0, 1]$, and α_1, α_2 are any two real numbers, then $\Lambda_\mu(\alpha_1 K_1 + \alpha_2 K_2) = \alpha_1 \Lambda_\mu(K_1) + \alpha_2 \Lambda_\mu(K_2)$.*
- 2) *Non-negativity: If $K \in C[0, 1]$ and $K \geq 0$, then $\Lambda_\mu(K) \geq 0$.*
- 3) *Continuity: If K_m is a non-increasing sequence (i.e., $K_1 \geq \dots \geq K_n \geq \dots$) of functions in L that converges to 0 for all x in $[0, 1]$, then $\Lambda_\mu(K_n) \rightarrow 0$.*

Proof. The first and the second properties are clear. Note that, when K is continuous, the integral $\int_0^1 K(t) d\mu(t)$ can be regarded as the Lebesgue integral for the measure m_μ defined as $m_\mu[a, b] = \mu(b) - \mu(a)$ and $m_\mu(1) = 0$. Therefore, the third property is an easy result of the dominated convergence theorem. \square

Let us now introduce the following Daniell integral (see Royden (1988) for instance) on the set, \mathbb{F} , of non-negative functions on $[0, 1]$: let $\{K_n\}_{n=1}^\infty$ be an arbitrary increasing sequence from $C[0, 1]$ converging pointwise to $K \in \mathbb{F}$, i.e., $K_n \uparrow K$. Then, the Daniell integral of K is defined as

$$D_\mu(K) = \lim_n \Lambda_\mu(K_n).$$

Daniell has shown that, given the three properties we mentioned in the previous lemma, this limit is independent of the choice of the sequence. It is important that

the value of the Daniell integral can be $+\infty$, if the limit is not bounded. There is also a Borel measure $\bar{\mu}$, on $[0, 1]$, given in intervals as

$$\bar{\mu}[a, b) = D_{\mu}(1_{[a, b)}) \text{ and } \bar{\mu}(\{1\}) = D_{\mu}(1_{\{1\}}),$$

where

$$D_{\mu}(K) = \int_0^1 K(t) d\bar{\mu}(t).$$

If the sequence $\int_0^1 K_n(t) d\mu(t)$ is bounded above, we say K is integrable. It is easy to see that $\bar{\mu}$ inherits all the properties of μ such as non-negativity, $\bar{\mu}[0, 1] = 1$, and $\text{supp}(\bar{\mu}) \subseteq [0, \epsilon_0]$. Now, let us consider some $K \in \mathbb{A}$ with a non-increasing sequence $\{K'_n\}_{n=1}^{\infty}$ from $C[0, 1]$ such that $K'_n \downarrow K$ on $[0, 1 - \frac{\epsilon_0}{2}]$ (we know that such sequence always exists given that K is bounded on $[0, 1 - \frac{\epsilon_0}{2}]$). Then, given that $\bar{\mu}$ is non-negative and $\text{supp}(\bar{\mu}) \subseteq [0, \epsilon_0] \forall \mu \in \Delta$, along with the pointwise continuity of ϱ , we have

$$\begin{aligned} \sup_{\mu \in \Delta} \int_0^1 K(t) d\bar{\mu}(t) &= \sup_{\mu \in \Delta} \int_0^{1-\epsilon_0} K(t) d\bar{\mu}(t) \\ &\leq \sup_{\mu \in \Delta} \int_0^{1-\epsilon_0} K'_n(t) d\bar{\mu}(t) \\ &= \sup_{\mu \in \Delta} \int_0^1 K'_n(t) d\bar{\mu}(t) \\ &= \Pi(K'_n) \downarrow \Pi(K). \end{aligned}$$

This inequality has two implications. First, K is Daniell integrable for all $\mu \in \Delta$. Second,

$$\sup_{\mu \in \Delta} \int_0^1 K(t) d\bar{\mu}(t) \leq \varrho(K). \quad (7)$$

On the other hand, let us assume that $\{K_n\}_{n=1}^{\infty}$ is a sequence in $C[0, 1]$ such that $K_n \uparrow K$ pointwise. Then we have

$$\sup_{\mu \in \Delta} \int_0^1 K(t) d\bar{\mu}(t) \geq \sup_{\mu \in \Delta} \int_0^1 K_n(t) d\bar{\mu}(t) = L^c(K_n) = \Pi(K_n) \uparrow \Pi(K).$$

This shows, on all members of \mathbb{A} , that

$$\sup_{\mu \in \Delta} \int_0^1 K(t) d\bar{\mu}(t) \geq \Pi(K). \quad (8)$$

Finally, (8) and (7) result in (2).

Now, let us assume that ϱ is co-monotone additive and introduce, for any continuous function K , the set $M(K)$ as follows:

$$M(K) = \left\{ \lambda \in \Delta \mid \int_0^1 K(t) d\lambda(t) = \sup_{\mu \in \Delta} \int_0^1 K(t) d\mu(t) \right\}.$$

Since Δ is non-empty, $M(K)$ is also non-empty. Now, we claim that for two non-decreasing continuous functions K_1 and K_2 , we have $M(K_1 + K_2) \subseteq M(K_1) \cap M(K_2)$. Let us take $\lambda \in M(K_1 + K_2)$; then by using L^c representation

$$\int_0^1 K_1(t) d\lambda(t) \leq \sup_{\mu \in \Delta} \int_0^1 K_1(t) d\mu(t), \quad (9)$$

$$\int_0^1 K_2(t) d\lambda(t) \leq \sup_{\mu \in \Delta} \int_0^1 K_2(t) d\mu(t) \quad (10)$$

and by co-monotonicity of ϱ , we have

$$\begin{aligned} \int_0^1 K_1(t) d\lambda(t) + \int_0^1 K_2(t) d\lambda(t) &= \int_0^1 (K_1(t) + K_2(t)) d\lambda(t) \\ &= \sup_{\mu \in \Delta} \int_0^1 (K_1(t) + K_2(t)) d\mu(t) \\ &= \Pi(K_1 + K_2) \\ &= \Pi(K_1) + \Pi(K_2). \end{aligned} \quad (11)$$

We can see that (9), (10), and (11) imply that (9) and (10) hold with equality, meaning that $\lambda \in M(K_1) \cap M(K_2)$. By induction, one can then infer $M(K_1 + K_2 + \dots + K_n) \subseteq M(K_1) \cap M(K_2) \cap \dots \cap M(K_n)$ for any n continuous and non-decreasing functions K_1, \dots, K_n . This means that, for every $n \in \mathbb{N}$, $M(K_1) \cap M(K_2) \cap \dots \cap M(K_n) \neq \emptyset$. By using the finite compact intersection lemma, one can deduce that

$$C := \bigcap \left\{ M(K) \mid \begin{array}{l} K \in C[0, 1] \\ K \text{ is non-decreasing} \end{array} \right\} \neq \emptyset.$$

Let us assume $\lambda \in C$. Then for any continuous and non-decreasing function K , we have $\Pi(K) = \int_0^1 K(t) d\lambda$. By using the same argument for introducing the Daniell integral above, one can show that λ can induce a measure on $[0, 1]$ such that $\Pi(K) = \int_0^1 K(t) d\lambda$, for any $K \in \mathbb{A}$. This completes the proof of Theorem 1. \square

Remark 2. It may be questioned why we did not adopt the same approach as in Ahmed et al. (2008) to introduce the following convex function

$$L^{cc}(K) = \begin{cases} \Pi(K) & K \in C[0, 1] \text{ and } K \text{ is non-decreasing} \\ +\infty & \text{o.w.} \end{cases},$$

and used instead its Fenchel-Moreau representation given by

$$L^{cc}(K) = \sup_{H \in \Delta_{cc}} \int_0^1 K(t) dH(t).$$

The matter of fact is, in this case (or any similar approach), to show that the set Δ_{cc} has non-negative members, we need to know that, for a continuous and non-decreasing function K , the sub-gradient $\partial L^{cc}(K)$ is non-empty. On the other hand, $\partial L^{cc}(K)$ is non-empty if K is in the interior of the domain:

$$\text{dom}(L^{cc}) = \{K \in C[0, 1] \text{ and } K \text{ is non-decreasing}\}.$$

However, it is not difficult to see that the interior of $\text{dom}(L^{cc})$ is empty.

The approach we have chosen above allows us to construct an appropriate reduction of Π to the set of non-decreasing continuous functions, which can be extended to the whole set $C[0, 1]$.

Remark 3. It is known that, for any non-empty set \mathbb{X} , the set of all real functions from \mathbb{X} to \mathbb{R} , endowed with the pointwise topology, is a topological vector space and each continuous functional f on this space can be described as $f(K) = \sum_{i=1}^n a_i K(x_i)$, for some $n \in \mathbb{N}$, $(a_i)_{i=1}^n \in \mathbb{R}^n$, and $(x_i)_{i=1}^n \in \mathbb{X}^n$ (see, for instance, Aliprantis and Burkinshaw (1978)). That is why one would guess, in the first place, that the same might hold for a weakly continuous co-monotone additive risk measure (which, as we have seen, does not).

Remark 4. Now let us compare our main result, Theorem 1, with a similar representation of law-invariant coherent risk measures by Kusuoka (2001). A coherent risk measure ϱ is a mapping from L^∞ to \mathbb{R} , with properties 1, 2, and 3 of Definition 1, and also with the following one:

$$4'. \text{ Sub-additivity: } \varrho(X + Y) \leq \varrho(X) + \varrho(Y), \forall X, Y \in L^\infty.$$

In Kusuoka (2001), it is shown that a law-invariant coherent risk measure ϱ that is $\sigma(L^\infty, L^1)$ lower semicontinuous can be represented as follows:

$$\varrho(X) = \sup_{m \in \mathcal{C}} \int_0^1 \text{CVaR}_\alpha(X) dm(\alpha),$$

where \mathcal{C} is a set of probability measures on $[0, 1]$ and

$$\text{CVaR}_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_t(X) dt, \forall X \in L^\infty.$$

However, this can be written as a double integral as follows:

$$\varrho(X) = \sup_{m \in \mathcal{C}} \int_0^1 \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_t(X) dt dm(\alpha).$$

By changing the variables, one gets

$$\begin{aligned} \int_0^1 \frac{1}{1-\alpha} \int_\alpha^1 \text{VaR}_t(X) dt dm(\alpha) &= \int_0^1 \int_0^t \frac{1}{1-\alpha} \text{VaR}_t(X) dm(\alpha) dt \\ &= \int_0^1 \text{VaR}_t(X) \left(\int_0^t \frac{1}{1-\alpha} dm(\alpha) \right) dt \\ &= \int_0^1 \text{VaR}_t(X) d\lambda_m(t), \end{aligned}$$

where $\lambda_m(t) = \int_0^t \int_0^s \left(\frac{1}{1-\alpha} dm(\alpha) \right) ds$. Note that λ_m is a non-decreasing function such that $\lambda_m(0) = 1 - \lambda_m(1) = 0$. Indeed, from the above, one can have

$$\lambda_m(1) = \int_0^1 \text{VaR}_t(X) \left(\int_0^t \frac{1}{1-\alpha} dm(\alpha) \right) dt \Big|_{X=1} = \int_0^1 \text{CVaR}_\alpha(X) dm(\alpha) \Big|_{X=1} = 1.$$

Therefore, by abusing the notation and using λ_m to denote a measure, $\lambda_m(a, b] = \lambda_m(b) - \lambda_m(a)$ can introduce a measure on $[0, 1]$. Finally, one can represent a law-invariant coherent risk measure in the following way:

$$\varrho(X) = \sup_{m \in \mathcal{C}} \int_0^1 \text{VaR}_\alpha(X) d\lambda_m(\alpha).$$

As one can see, this representation is very similar to what is expressed in Theorem 1, except that all members of \mathcal{C} have particular forms represented by $\lambda_m(t) = \int_0^t \int_0^s \left(\frac{1}{1-\alpha} dm(\alpha) \right) ds$.

References

Ahmed, S., D. Filipović, and G. Svindland (2008). A note on natural risk statistics. *Oper. Res. Lett.* 36(6), 662–664.

- Aliprantis, C. D. and O. Burkinshaw (1978). *Locally solid Riesz spaces*. New York: Academic Press [Harcourt Brace Jovanovich Publishers]. Pure and Applied Mathematics, Vol. 76.
- Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance* 9(3), 203–228.
- Assa, H. (2015a). On optimal reinsurance policy with distortion risk measures and premiums. *Insurance: Mathematics and Economics* 61(0), 70 – 75.
- Assa, H. (2015b). Risk management under a prudential policy. *Decisions in Economics and Finance*, 1–14.
- Assa, H. and K. M. Karai (2013). Hedging, Pareto optimality, and good deals. *J. Optim. Theory Appl.* 157(3), 900–917.
- Bertsimas, D., D. B. Brown, and C. Caramanis (2011). Theory and applications of robust optimization. *SIAM Review* 53(3), 464–501.
- Cai, J. and K. S. Tan (2007). Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures. *Astin Bull.* 37(1), 93–112.
- Cai, J., K. S. Tan, C. Weng, and Y. Zhang (2008). Optimal reinsurance under VaR and CTE risk measures. *Insurance Math. Econom.* 43(1), 185–196.
- Delbaen, F. (2002). Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pp. 1–37. Berlin: Springer.
- Denneberg, D. (1994). *Non-additive measure and integral*, Volume 27 of *Theory and Decision Library. Series B: Mathematical and Statistical Methods*. Kluwer Academic Publishers Group, Dordrecht.
- Ekeland, I. and R. Témam (1999). *Convex Analysis and Variational Problems*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.
- Föllmer, H. and A. Schied (2002). Robust preferences and convex measures of risk. In *Advances in finance and stochastics*, pp. 39–56. Berlin: Springer.
- Frittelli, M. and E. Rosazza Gianin (2002). Putting order in risk measures. *Journal of Banking & Finance* 26(7), 1473 – 1486.
- Greco, G. H. (1982). Sulla rappresentazione di funzionali mediante integrali. *Rendiconti del Seminario Matematico della Università di Padova* 66, 21–42.

- Grothendieck, A. (1973). *Topological vector spaces*. New York: Gordon and Breach Science Publishers. Translated from the French by Orlando Chaljub, Notes on Mathematics and its Applications.
- Kou, S., X. Peng, and C. C. Heyde (2013). External risk measures and basel accords. *Math. Oper. Res.* 38(3), 393–417.
- Kusuoka, S. (2001). On law invariant coherent risk measures. In *Advances in mathematical economics, Vol. 3*, Volume 3 of *Adv. Math. Econ.*, pp. 83–95. Tokyo: Springer.
- Nakano, Y. (2004). Efficient hedging with coherent risk measure. *J. Math. Anal. Appl.* 293(1), 345–354.
- Quaranta, A. G. and A. Zaffaroni (2008). Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking & Finance* 32(10), 2046 – 2056.
- Rockafellar, R. T. (1997). *Convex analysis*. Princeton Landmarks in Mathematics. Princeton, NJ: Princeton University Press. Reprint of the 1970 original, Princeton Paperbacks.
- Royden, H. (1988). *Real Analysis*. Mathematics and statistics. Macmillan.
- Rudin, W. (1987). *Real and Complex Analysis, 3rd Ed.* New York, NY, USA: McGraw-Hill, Inc.
- Rudin, W. (1991). *Functional analysis* (Second ed.). International Series in Pure and Applied Mathematics. New York: McGraw-Hill Inc.
- Wang, S. S., V. R. Young, and H. H. Panjer (1997). Axiomatic characterization of insurance prices. *Insurance Math. Econom.* 21(2), 173–183.
- Xing, X., J. Hu, and Y. Yang (2014). Robust minimum variance portfolio with l-infinity constraints. *Journal of Banking & Finance* 46, 107 – 117.
- Zymler, S., B. Rustem, and D. Kuhn (2011). Robust portfolio optimization with derivative insurance guarantees. *European Journal of Operational Research* 210(2), 410 – 424.